Unmasking All Forms of Cancer: Toward Integrated Maps of All Tumor Subtypes

Distinguished Lecture in Causal Discovery
Center for Causal Discovery
(U. Pitt, Carnegie Mellon, Pitt. SCC, Yale)
University of Pittsburgh, PA. Feb 16, 2017

Josh Stuart, Professor
Baskin Engineering Endowed Chair
UC Santa Cruz
THE -$1 GENOME IS HERE

Circa 2014
“Is the $1000 Genome real?”

Plateau?

Sources: NIH: www.genome.gov/sequencingcosts; UC San Diego, 1/14/14: Illumina breaks genome cost barrier
THE -$1 GENOME IS HERE

Further Advances

Cost per Genome

Sources: NIH: www.genome.gov/sequencingcosts; UC San Diego, 1/14/14: Illumina breaks genome cost barrier
THE -$1 GENOME IS HERE

Your Genome Costs Less Than Your Phone

Cost per Genome

Sources: NIH: www.genome.gov/sequencingcosts; UC San Diego, 1/14/14: Illumina breaks genome cost barrier
THE -$1 GENOME IS HERE

Companies will Pay.

Interpretation is where Value Is.

Negative Dollar Genome

Cost per Genome

Sources: NIH: www.genome.gov/sequencingcosts; UC San Diego, 1/14/14: Illumina breaks genome cost barrier
THE POTENTIAL FOR DNA AND COMPUTING TO TRANSFORM MEDICINE IS NOT BEING REALIZED

Opportunities to save lives are lost every day

UC Santa Cruz Genomics Institute
CANCER GENOMICS: A VIEW INSIDE TUMOR CELLS

Normal Cell

Tumor

mutation
SEQUENCE THE CANCER GENOME

Germline DNA from blood

Tumor DNA

Billions of short DNA reads

UC Santa Cruz Genomics Institute
Cancer Data Revolution

Genome Data Commons

- 1000 Genomes Project: 220 TB
- National Human Genome Research Institute: 68 TB
- Alzheimer's Disease Sequencing Project: 34 TB
- National Heart, Lung, and Blood Institute: 32 TB
- Human Microbiome Project: 31 TB
- ENCODE Project: 25 TB
- ARRA Autism Sequencing Collaboration: 20 TB
- NHGRI Large-Scale Sequencing Program: >5,000 TB

UC Santa Cruz Genomics Institute
Sequenced Cancer. Now What?
Sequenced Cancer. Now What?

- Interpret DNA changes w/ functional information
- Transcriptome key to state read-out
- Connect-the-dots with pathway inference
PERSONALIZED NETWORKS FOR TARGETING

Patient DTB-011

- BCL2 – B-cell lymphoma related
  - Blocks apoptosis of cells.
  - Targeting in PCa (Zielinski Cancer J 2013)
- GSK38 – glycogen synthase kinase 3
  - Inhibitors reduce PCa growth (Darrington Int J Cancer 2012).
- MAPK8 (aka JUN Kinase)
  - siRNA induces apoptosis in PCa (Parra Int J Mol Med 2012)
- MAPK14 (aka p38)
  - Inhibitors may promote mets
- HRAS
  - Synthetic lethal with JNK (above) (Zhu Genes Cancer 2010)
- SHC1 – Src homolog
  - ERK and TGFB signaling
PERSONALIZED NETWORKS FOR TARGETING

Patient DTB-011

- BCL2 – B-cell lymphoma related
  - Blocks apoptosis of cells.
  - Targeting in PCa (Zielinski Cancer J 2013)
- GSK38 – glycogen synthase kinase 3
  - Inhibitors reduce PCa growth (Darrington Int J Cancer 2012).
- MAPK8 (aka JUN Kinase)
  - siRNA induces apoptosis in PCa (Parra Int J Mol Med 2012)
- MAPK14 (aka p38)
  - Inhibitors may promote mets
- HRAS
  - Synthetic lethal w/ JNK (above) (Zhu Genes Cancer 2010)
- SHC1 – Src homolog
  - ERK and TGFB signaling

Patient 11-specific Drug Combinations
Outline:
Interpreting A Cancer Genome (N-of-1)

- Identify the closest known form
- Tailor the pathway model to fit an individual tumor’s unique combination of events
Outline:
Interpreting A Cancer Genome (N-of-1)

- Identify the closest known form
- Tailor the pathway model to fit an individual tumor’s unique combination of events
Identify all the forms of cancer?
Identify all the forms of cancer?
Identify all the forms of cancer?

Responsive Subtypes To Treatment

Oncogenic Process

Pathway

Treatment Outcome

Response

No Response

Cell-Of-Origin
6 Data Platforms – Subtypes from each

mRNA (Hoadley, UNC)

microRNA (Robertson, UBC)

Protein (Akbani, MDACC)

DNA Copy Number (Cherniack, Broad)

DNA Methylation (Shen, USC)

Exome-Mutations (not used) (Uzunangelov, UCSC)
MULTIPLE TYPES OF GENOMICS DATA

- Expression
- DNA Methylation
- Structural Variation
- Exome Sequences
- Copy Number Alterations

$2^n$ combos
FLOOD OF DATA ANALYSIS CHALLENGES

Genomics, Functional Genomics, Metabolomics, Epigenomics =

This is What it Does to You

Multiple, Possibly Conflicting Signals

Expression

Copy Number Alterations

Chromosome
WHAT GUIDES DO WE HAVE TO INFER THE LAWS GOVERNING INTERPLAY OF CELLULAR SYSTEMS?

Relationships of the motions confusing!
What guides do we have to infer the laws governing interplay of cellular systems?

Model = Simpler Explanation
Curated and/or Collected
Reactome
KEGG
Biocarta
NCI-PID
Pathway Commons
Pathway Recognition Algorithm Using Data Integration on Genomic Models (PARADIGM)

Charlie Vaske, Steve Benz
PARADIGM
Gene Model to Integrate Data

3-state discrete variables

relative to non-cancer, is this sample: up, same, down?

Charlie Vaske, Steve Benz
PARADIGM Gene-level Model

3-state discrete variables

relative to non-cancer, is this sample: up, same, down?

<table>
<thead>
<tr>
<th>CNA \ Exp</th>
<th>Down</th>
<th>Same</th>
<th>Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down</td>
<td>0.90</td>
<td>0.09</td>
<td>0.01</td>
</tr>
<tr>
<td>Same</td>
<td>0.05</td>
<td>0.90</td>
<td>0.05</td>
</tr>
<tr>
<td>Up</td>
<td>0.01</td>
<td>0.09</td>
<td>0.90</td>
</tr>
</tbody>
</table>
PARADIGM Gene Model to Integrate Data

- GeneCopy Number
- Expression State
- Protein Level
- Protein Activity

Array CGH, SNP chips
Transcriptomics

Variable
- Factor - interaction term

Charlie Vaske, Steve Benz
PARADIGM Gene Model to Integrate Data

GeneCopy Number

Expression State

Protein Level

Protein Activity

Array CGH, SNP chips

Transcriptomics

Proteomics, mutations

Proteomics, mutations

Variable

- Factor - interaction term

Charlie Vaske, Steve Benz
PARDIGM Gene Model to Integrate Data
Interactions Matter

➢ Given information about the expression of TP53 alone
➢ Reasoning predicts apoptosis is in tact in these cells.

Charlie Vaske, Steve Benz
Interactions Matter

Given the interaction and data about MDM2.

 ➢ apoptosis inference reversed

Quantitative Output

Log likelihood Ratio:

\[
\log \left( \frac{P(Data|Apoptosis_{active})}{P(Data|Apoptosis_{inactive})} \right) = \log \left( \frac{P(Data|Apoptosis_{active})}{P(Data|Apoptosis_{active})} \right) - \log \left( \frac{P(Apoptosis_{active})}{P(Apoptosis_{inactive})} \right)
\]

log odds of state and data

prior log odds

Charlie Vaske, Steve Benz
CELL CIRCUITRY – BAD FOR HUMAN CONSUMPTION

Circuitry (Hairball)

What it Does to You
CELL CIRCUITRY – GREAT FOR COMPUTER CONSUMPTION

Circuitry (Hairball) + Data
COMPUTING! (AWFUL FOR HUMAN CONSUMPTION!)

Circuitry (Hairball) + Data = Insights
INTEGRATED MAP TO RULE THEM ALL

Expression
Exome Sequences
DNA Methylation
Copy Number Alterations
Structural Variation
INTEGRATED MAP TO RULE THEM ALL

Patient Samples (3491)

Pathway Concepts (13,480)

Copy Num Alteration

Hoadley et al Cell 2014
• ~90% of samples cluster with their tissue
Viewing Gene Programs on the TumorMap

ER Signaling “Weather Map”

KIRC Show Moderate ER signaling

BRCA Luminals Show High ER signaling

Denise Wolf, UCSF
Are disease-specific AWG subtypes recapped in TumorMap?

- Good agreement overall.

Pattern on another molecular map adds insight.

Newton, Baertsch, UCSC
• BLCA diverge into bladder-enriched, squamous, and LUAD-enriched islands

Hoadley, Cell, Aug 2014
INTEGRATED SUBTYPING OF BLCA DISTINGUISHES PATIENT OUTCOMES

- COCA clusters distinguish different survival classes for BLCA

Hoadley, Cell, Aug 2014
PanCan-33: PanCanAtlas

- **33 Tumor Types**
- **11,053 Total Cases**
- **Latest Publication Restrictions Lift in December, 2015 (e.g. testicular)**
- **Average cases: 335**
- **Median cases: 308**
- **BRCA most cases: 1100**
- **CHOL least cases: 36**
PanCan-33 TumorMap

Colors show Tissue of origin.

Newton, Baertsch, UCSC
Integrated map reveals pancan subtypes

PanCan Subtype (not seen in original analysis)
Integrated map reveals new subtypes

What characterizes these tumors?
Enriched for t/B and IFN immune (D. Wolf’s) programs
BLCA divergence in Pan-Can-33

- BLCA diverge into several more subtypes
PANCAN-12
RECLASSIFICATION RATE = 1 in 10
PANCAN-33
RECLASSIFICATION RATE = 1 in 5
PANCAN FOR N=1 PATIENTS

UC Santa Cruz Genomics Institute

Genomic mapping

PanCan-33 Map:

Lung Cancer  Bladder Cancer

Genomic mapping
Clinical Genomics Trials
-- UCSF, PNOC (15 pts)
-- UCI, CHOC (40 pts)
-- Stanford (100pts)

Genomic characterization data; Clinical data

Clinical leads

Outcome measures:
• New clinical leads
• New evidence for clinical leads
• New/refined molecular diagnoses

Large adult genomic databases (TCGA, ICGC, SU2C)

TumorMap
Xena
PrecisionImmuno
NuMedii

MedBook

Treehouse pediatric cancer Data (including TARGET)
WHERE DO Childhood Samples MAP?

Olena Morozova
Yulia Newton
Analysis of POG samples in the context of other cancers

- **TH005-PED3**
  - Clusters with Pheochromocytoma and Paraganglioma (pancan30) and with Neuroblastoma (pancan14)

- **TH002_NBL**
  - Clusters with Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (pancan30) and with Neuroblastoma (pancan14)

- **TH004_SCC**
  - Clusters with Head and Neck Squamous Cell Carcinoma

- **TH006_NBL**
  - Clusters with Pheochromocytoma and Paraganglioma (pancan30) and with Neuroblastoma (pancan14)

- **TH007_NF**
  - Clusters with Breast Invasive Carcinoma

- **TH003_NBL**
  - Clusters with Neuroblastoma

- **TH001_SARC**
  - Clusters with Neuroblastoma ALK fusion tumors
Observation: TH001 pediatric sarcoma groups with neuroblastoma ALK-mutant samples.
ALK POTENTIAL TARGET FOR PATIENT 1 BASED ON PAN-CANCER ANALYSIS

Normalized relative ALK expression level

- EML4-ALK lung (N=2)
- ALK-amp neuroblastoma cohort (N=15)
- TH001 sarcoma
- Non-ALK neuroblastoma cohort (N=270)
- Sarcoma cohort (N=172)
TWO NEW TREATMENTS FOR PATIENT 1

Cancer cell:
- **ALK**
- **FGFR1**
- **IL6R**

**JAK1**

Uncontrolled cell growth

Normal cell:
- **ALK**
- **FGFR1**
- **IL6R**

**JAK1**

Controlled cell growth
California Kids Cancer Comparison (CKCC)

Childhood cancer in California

1800 children are diagnosed with cancer each year in California.

400-500 of them will not respond to conventional treatment, and would die within 5 years.

100-150 of those patients are recruited for clinical trials.

Current clinical trials:
Repurpose existing drugs matched to specific genomic defects in each tumor

Expected: new treatment options for ~10% of kids

Analysis of genomic information from each tumor

Genomic defect with match to existing drug identified?

yes

no

Expected: ~90% of patients will still die

CKCC GOAL:
More kids helped since more matches lead to more new treatment options

More repurposed therapies identified based on tumor comparisons

State of the art computational approach to find more matches to existing drugs

Big data technology:
Integrate all data for each tumor and compare each tumor's genomic information with that of 1000s of other tumors.

BECAUSE OF PATIENT 1...
WHAT WE ARE DOING NOW: MOLECULAR DETECTIVES

- Current cases of children with cancer
- TH008: 2-year-old diagnosed with Stage 4 Hepatoblastoma (liver cancer)
- Underwent two chemo protocols and two surgeries
- In need of new treatment options
- Foundation Medicine test revealed CTNNB1G34V mutation
TH008 IS MORE SIMILAR TO ADULT LIVER TUMORS THAN EXPECTED

Bird’s eye view (tumors colored by disease)

Zoom in on the patient (tumors colored by disease)

Hepatocellular carcinoma (liver)

Colangiocarcinoma
<table>
<thead>
<tr>
<th>Target</th>
<th>Drug</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aurora kinases</td>
<td>Pazopanib</td>
<td>Clinical trial</td>
</tr>
<tr>
<td>IGF1R</td>
<td>Metformin</td>
<td>Off-label</td>
</tr>
<tr>
<td>ABCC2</td>
<td>Simvastatin plus chemo</td>
<td>Clinical trial</td>
</tr>
<tr>
<td>JAK/STAT</td>
<td>Ruxolitinib</td>
<td>Off-label</td>
</tr>
</tbody>
</table>

Turns out trial of pazopanib is opening up at Stanford and so treating oncologist chose this option
Outline:
Interpreting A Cancer Genome (N-of-1)

➢ Identify the closest known form

➢ Tailor the pathway model to fit an individual tumor’s unique combination of events
PERSONALIZED NETWORKS FOR TARGETING

Patient DTB-011

Linking Network

- **BCL2** – B-cell lymphoma related
  - Blocks apoptosis of cells.
  - Targeting in PCa (Zielinski *Cancer J* 2013)
- **GSK3B** – glycogen synthase kinase 3
  - Inhibitors reduce PCa growth (Darrington *Int J Cancer* 2012).
- **MAPK8** (aka JUN Kinase)
  - siRNA induces apoptosis in PCa (Parra *Int J Mol Med* 2012)
- **MAPK14** (aka p38)
  - Inhibitors may promote mets
- **HRAS**
  - Synthetic lethal with JNK (above) (Zhu *Genes Cancer* 2010)
- **SHC1** – Src homolog
  - ERK and TGFB signaling
PERSONALIZED NETWORKS FOR TARGETING

**Patient DTB-011**

**Linking Network**

- **BCL2** – B-cell lymphoma related
  - Blocks apoptosis of cells.
  - Targeting in PCa (Zielinski *Cancer J* 2013)

- **GSK3B** – glycogen synthase kinase 3
  - Inhibitors reduce PCa growth (Darrington *Int J Cancer* 2012).

- **MAPK8** (aka JUN Kinase)
  - siRNA induces apoptosis in PCa (Parra *Int J Mol Med* 2012)

- **MAPK14** (aka p38)
  - Inhibitors may promote mets

- **HRAS**
  - Synthetic lethal w/ JNK (above) (Zhu *Genes Cancer* 2010)

- **SHC1** – Src homolog
  - ERK and TGFB signaling

**Patient 11-specific Drug Combinations**
ASIDE: WHAT ARE THE IMPORTANT “EVENTS” IN A TUMOR?

• Lots of Copy number, point mutations

• Which are passengers? Which drivers?

• What does data reveal about essential signaling?

• Aside: Just identifying variants is hard!
A needle in a human genome haystack

- A human genome has **23** chromosomes.
- **6 billion** individual DNA basepairs per genome.
- A **single basepair** error can be a disease mutation.
Distinguish True Variation from Artifact

SNV

sequencing errors
Mutation Callers Give Different Answers ...

**SNVs**

- **Baylor**: 1,245
- **12,036**: 2,474
- **1,495**: 14,320
- **UCSC**: 1,114
- **14,122**:

**SVs**

- **BreakDancer**: 14193
- **GASV**: 4103
- **CREST**: 3190
- **1005**: 871
- **1122**:

---

**TCGA Network mutation calls**

Renal Clear Carcinoma: Somatic calls from 297 Samples
DREAM for the best method(s)

Crowd-source for best mutation detectors.

Define dataset and goal.

Put out incentives (talks, papers, $$)

Collaboration: OICR, TCGA, UCSC, SAGE
Results of DREAM-SMC

Participation At Closing Time:
- 345 contestants
- 948 entries on 4 *in silico* genomes

On-going post-challenge submissions (*living benchmark*)

Key insights into simulating cancer genomes (*BamSurgeon*)
Wisdom of the Crowds for DREAM-SMC

Accuracy (F-score)

Ensemble of top $k$ methods

Individual methods

Accuracy of single best method

Ave of all methods matches best single

Ewing et al. Nat Meth 2014
Negative Results Reveal False-Positive Signature

Many methods see “ghost” C->T mutations.

Matches a signature reported in a high-profile paper...

Ewing et al. Nat Meth 2014
ASIDE: WHAT ARE THE IMPORTANT "EVENTS" IN A TUMOR?

• No current consensus on how to interpret variants.

• There are many algorithms and boutique bakeoffs

Tokheim et al PNAS 2016
PATHWAY REASONING

High Inferred Activity
Inference using all neighbors

Inference using downstream neighbors

Inference using upstream neighbors

Low Inferred Activity
mutated gene

Sam Ng, Bioinformatics 2012
RB1 LOF (GBM)
RB1 LOF (GBM)
RB1 LOF (GBM)

Inferred Upstream Expression RB1 Mutation
RB1 LOF (GBM)
RB1 LOF (GBM)
Upstream and Downstream Genes

PARADIGM Expression Mutation Status of focus gene (RB1)

RB1 LOF (GBM)
Upstream and Downstream Genes

PARADIGM Expression
Mutation Status of focus gene (RB1)

RB1 LOF (GBM)

High Activator Activity
Upstream and Downstream Genes

PARADIGM
Expression
Mutation Status of focus gene (RB1)

Low Inhibitor Activity

RB1 LOF (GBM)
Gain-of-Function (LUSC)

- P-Shift Score
- PARADIGM downstream
- PARADIGM upstream
- Expression
- Mutation

NFE2L2

Sam Ng
PARADIGM-Shift gives orthogonal view of the importance of mutations (LUSC)

➢ Enables probing into infrequent events
➢ Can detect non-coding mutation impact (pseudo FPs)
➢ Can detect presence of pathway compensation for those seemingly functional mutations (pseudo FPs)
➢ Extend beyond mutations

Limited to genes w/ pathway representation

HIF3A (n=7)
TBC1D4 (n=9) (AKT signaling)
NFE2L2 (29)
MAP2K6 (n=5)
MET (n=7) (gefitinib resistance)
GLI2 (n=10) (SHH signaling)
CDKN2A (n=30)
EIF4G1 (n=20)
AR (n=8)
PERSONALIZED NETWORKS FOR TARGETING

**Patient**

**Mutations**

- TP53
- ERBB2
- NOTCH1

**Signature**

- RNA-seq data informs a set of genes are significantly up- and another down-regulated.

- Match profile with a known cancer subtype to obtain robustness of transcripome classification.
• Link mutations to transcriptional changes with heat-diffusion on networks (e.g. PPI or curated).

Signature

Signature Word Cloud

Summary

- Cytokine-cytokine receptor interaction
- Apoptosis
- Cell cycle
- Glucose metabolism
PERSONALIZED NETWORKS FOR TARGETING

Patient

DTB-011

Mutations

TP53

ERBB2

NOTCH1

Signature

See Master Regulator Analysis (Califano Lab)

Infer Active Transcription Factors

RNA-Seq

RSEM

TF Targets

Activation Score

MARINα
PERSONALIZED NETWORKS FOR TARGETING

Infer Active Transcription Factors

TFs: Inferred Transcription Factors

Signature

TF’s targets have low expression
TIEDE: LINKING MUTATIONS TO SIGNATURES

• Still need connections between mutations and inferred TFs
PERSONALIZED NETWORKS FOR TARGETING

- Still need connections between mutations and inferred TFs

Patient
DTB-011
Mutations
TP53
ERBB2
NOTCH1

TFs: Inferred Transcripti
RPA1
CREB1
ATF2
MYB

Signature
KLK2
TAGLN
DICER1
NDRG1

• Still need connections between mutations and inferred TFs
PERSONALIZED NETWORKS FOR TARGETING

Mutations

Signature

Inferred Transcript

TP53

ERBB2

NOTCH1

Patient

DTB-011

Linking Network

Heat diffusion approaches

“Sources”

“Targets”

Background Network

Time = 0

Time = 1

e.g. Bader 2010, Vandin 2012, Paull 2013, Hofree 2013)
Mets show a distinct phosphorylation pattern, when compared with treatment-naive samples. In total, 8,051 peptides were measured.

Question: Does a network solution using mutations and TFs include the activated kinases detected by protein Mass-Spec?
TieDIE Networks Embed Activated Proteins

Are Linkers More Activated?

Drake, Paull et al Cell 2016
TieDIE Networks Embed Activated Proteins

Are Linkers More Activated?

![Differential Activity Graph]

$p < 4.5e-6$ (KS)

Drake, Paull et al Cell 2016
Master Regulator Analysis (MRA) on *Phosphoproteomic* data

Classic MRA: target gene expression -> protein activity

**Proteomic MRA**: kinase target phosphorylation -> protein activity

*Chen et al., Califano 2014*

Drake, Paull et al *Cell* 2016
Master Regulator Analysis on **Phosphoproteomic** data

MAPK14, PRKDC, CDK1, AKT1, SRC, PRKAA2....

*Plot made with VIPER Bioconductor R package*

source("https://bioconductor.org/biocLite.R")
biocLite("viper")

Drake, Paull et al *Cell* 2016
TieDIE Networks Embed Activated Proteins

Are Linkers More Activated?  
Are ~Active TFs near ~Active Kinases?

p < 4.5e-6 (KS)

Drake, Paull et al. Cell 2016
TieDIE Networks Embed Activated Proteins

Are Linkers More Activated?

Are ~Active TFs near ~Active Kinases?

Drake, Paull et al Cell 2016
Scaffold network for CRPC from eclectic data

Gene Expression
88 samples

Phosphopeptides
27 samples

Data Platforms

Differential Analysis
Metastatic CRPC:Rx Naïve PrCa

Gene Expression Master Regulators

Tyrosine Peptide Data (Kinases)
Drake et al. (2013)

Inferred Kinase Master Regulators

Gene Expression Regulators

Inferred Kinase Activity

Phosphorylated Kinases (from MS data)

74 0 11
0 0 3
21

Drake, Paull et al Cell 2016
Scaffold network for metastatic prostate from diverse data

Drake, Paull et al
Cell 2016

(1) Scaffold Network

Data Platforms

Gene Expression
88 samples

Phosphopeptides
27 samples

Differential Analysis
Metastatic CRPC:Rx Naive PrCa

Gene Expression Master Regulators
Tyrosine Peptide Data (Kinases)
Drake et al. (2013)
Inferred Kinase Master Regulators

Gene Expression Regulators
Inferred Kinase Activity

Phosphorylated Kinases (from MS data)

(2) TiedIE Analysis

Mutation Data (TCGA & CRPC)

Physical Interaction Pathway Database (Multinet)

Drake, Paull et al Cell 2016
PERSONALIZED NETWORKS FOR TARGETING

Mutations

Inferred Transcriptomic Factors

Signature Genes

Patient DTB-011

Linking Network

Time = 0
Diffuse Heats

Time = 1

Signature Genes

Linking Genes

Patient

[Diagram showing network of genes and proteins with interactions regulated by time and diffusion processes]
N-of-1 Patient-specific Network Approach Overview

Drake, Paull et al Cell 2016
N-of-1 Patient-specific Network Approach Overview

Patient-Specific Mutations

TP53, ERBB2, NOTCH1

Pathway Network Scaffold

Patient-Specific Gene Expression

Data Platforms

VIPER Algorithm

TieDIE$^1$

ID Master Regulators (ala Califano)

TieDIE$^2$

Patient-Specific Network Model

Drake, Paull et al Cell 2016
Patient RA40

Drake, Paull et al Cell 2016
Patient RA40

Drake, Paull et al. Cell 2016
Patient
RA40

Drake, Paull et al Cell 2016
Patient RA40

Drake, Paull et al Cell 2016
Patient RA40

Drake, Paull et al Cell 2016
PI3K-AKT-mTOR pathway

Drake, Paull et al Cell 2016
Network-based selection of targets and target combinations for individual patients
Network-based selection of targets and target combinations for individual patients

Patient 40

Drake, Paull et al. Cell 2016
Network-based selection of targets and target combinations for individual patients

Patient 40

Drake, Paull et al Cell 2016
Network-based selection of targets and target combinations for individual patients

Drake, Paull et al Cell 2016

Patient 40
Treat with AKT1 inhibitor
Network-based selection of targets and target combinations for individual patients

Drake, Paull et al. Cell 2016
Network-based selection of targets and target combinations for individual patients

Patient 30

Drake, Paull et al Cell 2016
Network-based selection of targets and target combinations for individual patients

Patient 30

Drake, Paull et al Cell 2016
Network-based selection of targets and target combinations for individual patients

Patient 30

Treat with AKT1 & SRC inhibitor

Drake, Paull et al Cell 2016
TAKE-HOME MESSAGES

• Pan-Cancer analysis reveals strong tissue-of-origin signals.
  • But ~10% reclassified associated w/ survival.

• Adult signatures can inform novel pan-cancer connections for treatment avenues in pediatric cancer

• Integration of proteomic data with other ‘omics’ data reveals signaling pathways in metastatic prostate cancer.

• Patient-specific hierarchy of clinically actionable pathways for therapy.
Future Directions

- Integrative methods for variant interpretation
- Pathway ID for sub-clones & stroma & immune, etc
- Formal causal models to reveal pathway “weaknesses”
- Single cell (e.g. cfDNA) pathway analysis for early detection
UCSC Integrative Genomics Group

Evan Paull
Vlado Uzunangelov
Kiley
Yulia Newton
David Haan
Artem Sokolov
James Durbin
Robert Baertsch
Chris Wong
Acknowledgments

UCSC Genome Browser Staff
- David Haussler
- Mark Diekkins
- Melissa Cline
- Jorge Garcia
- Erich Weiler

UCSF / Buck Institute for Aging
- Chris Benz, Buck
- Christina Yau, Buck
- Denise Wolf, UCSF
- Laura van’t Veer, UCSF
- Eric Collisson, UCSF

Collaborators
- Chuck Perou, UNC
- Katie Hoadley, UNC

Witte Lab
- Justin Drake, (now at Rutgers)
- Owen Witte, HHMI

UCSC Medbook Team
- Ted Goldstein

UCSC Cancer Genomics
- Jing Zhu
- Sofie Salama
- Teresa Swatlowski
- Brian Craft

UCSC Tree House Project
- Olena Morozova
- Melissa Cline

OSRACLE PCF LINCS

THE CANCER GENOME ATLAS